

Large Language Models are Zero-Shot Rankers for Recommender Systems

Year: 2023 | Citations: 438 | Authors: Yupeng Hou, Junjie Zhang, Zihan Lin

Abstract

Recently, large language models (LLMs) (e.g., GPT-4) have demonstrated impressive general-purpose task-solving abilities, including the potential to approach recommendation tasks. Along this line of research, this work aims to investigate the capacity of LLMs that act as the ranking model for recommender systems. We first formalize the recommendation problem as a conditional ranking task, considering sequential interaction histories as conditions and the items retrieved by other candidate generation models as candidates. To solve the ranking task by LLMs, we carefully design the prompting template and conduct extensive experiments on two widely-used datasets. We show that LLMs have promising zero-shot ranking abilities but (1) struggle to perceive the order of historical interactions, and (2) can be biased by popularity or item positions in the prompts. We demonstrate that these issues can be alleviated using specially designed prompting and bootstrapping strategies. Equipped with these insights, zero-shot LLMs can even challenge conventional recommendation models when ranking candidates are retrieved by multiple candidate generators. The code and processed datasets are available at <https://github.com/RUCAIBox/LLMRank>.